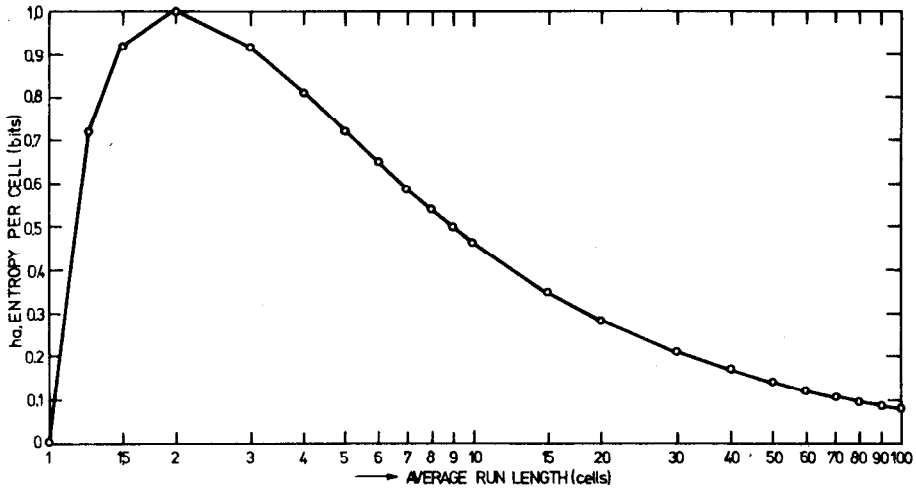

Fig. 1. Entropy per cell as function of average run length.

distributed:

$$p(t_1) = \begin{cases} be^{-bt_1}, & \text{for } t_1 \geq 0 \\ 0, & \text{for } t_1 < 0 \end{cases} \tag{19}$$

where t_1 is the (continuous) run length. The average run length is

$$\bar{t}_1 = b. \tag{20}$$

Now, suppose the continuous run length t_1 is quantized to obtain the discrete run length t:

$$t = [t_1] + 1 \tag{21}$$

where

$$[t_1] = \text{largest integer in } t_1. \tag{22}$$

Then, t takes on the integer values $1,2,3,\cdots$, and

$$\begin{aligned} q_i &= \Pr\{t = i\} \\ &= e^{-b(i-1)} - e^{-bi} = (e^b - 1)e^{-bi}. \end{aligned} \tag{23}$$

According to (23)

$$\bar{t} = \frac{1}{1 - e^{-b}}. \tag{24}$$

Letting $\bar{t} = a$, we obtain

$$a = \frac{1}{1 - e^{-b}} \tag{25}$$

or

$$b = \log \frac{a}{a - 1}. \tag{26}$$

Substituting (26) into (23), we obtain

$$q_i = \frac{1}{a - 1} e^{-[\log a - \log (a-1)]i} \tag{27}$$

which is identical to (15). Therefore, the quantized Poisson square wave achieves the maximum entropy given by (16).

IV. Relation to Capon's Markov Chain Model

After the aforementioned analysis was performed, it was found that the quantized Poisson square wave is identical to Capon's first-order Markov chain model [1], if we set

$$P(0 \mid 0) = P(1 \mid 1) = \frac{a - 1}{a}. \tag{28}$$

Therefore, although Capon apparently did not realize it, the saving in bits predicted by his model is actually a lower bound for any two-level source with average white run length $1/(1 - P(1 \mid 1))$ and average black run length $1/(1 - P(0 \mid 0))$, because the run lengths in Capon's model are independent, and the exponential distributions of the white run lengths and the black run lengths ensure that both achieve the maximum entropy.

References

[1] J. Capon, "A probabilistic model for run-length coding of pictures," *IRE Trans. Inform. Theory*, vol. IT-5, pp. 157–163, Dec. 1959.

A Low-Rate Improvement on the Elias Bound

LLOYD R. WELCH, ROBERT J. McELIECE, MEMBER, IEEE, AND HOWARD RUMSEY, JR.

***Abstract*—An upper bound on the minimum distance of binary blocks codes, which is superior to Elias' bound for $R < 0.0509^+$, is obtained. The new bound has the same derivative $(-\infty)$ at $R = 0$ as Gilbert's lower bound. (Elias' bound has derivative $-\ln 2$ at $R = 0$).**

I. Introduction

For R between 0 and 1 denote by $d(n,R)$ the largest possible minimum distance for a binary block code of length n and rate $\geq R$. It is unknown whether

$$D(R) = \lim_{n \to \infty} \frac{1}{n} d(n,R)$$

exists, so let us define

$$\bar{D}(R) = \limsup_{n \to \infty} \frac{1}{n} d(n,R)$$

$$\underline{D}(R) = \liminf_{n \to \infty} \frac{1}{n} d(n,R).$$

Until now the best bounds on $\bar{D}(R)$ and $\underline{D}(R)$ have been

$$\underline{D}(R) \geq f(1 - R) \qquad \text{(Gilbert, 1952)}$$

$$\bar{D}(R) \leq 2f(1 - R) - 2f^2(1 - R) \qquad \text{(Elias, 1960)}$$

Manuscript received January 21, 1974. This work was supported by NASA under Contract NAS 7-100.

The authors are with the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, Calif. 91103.

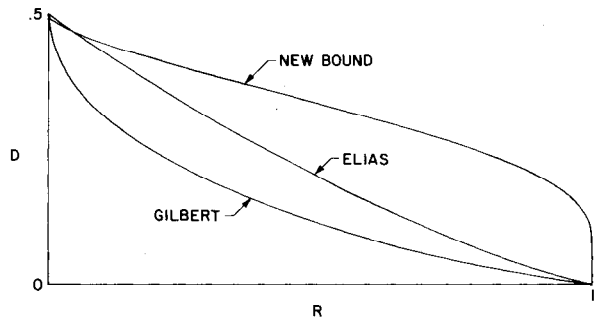

Fig. 1. New bound and old bounds.

where f is the inverse of the binary entropy function, i.e., $f(x) = y$, if and only if $0 \le y \le \frac{1}{2}$ and $H_2(y) = -y \log_2 y - (1 - y) \cdot \log_2 (1 - y) = x$. (For proofs, see [1, ch. 13]). The Gilbert and Elias bounds agree at $R = 0$ and $R = 1$, establishing the existence of the limits $D(0) = \frac{1}{2}$ and $D(1) = 0$, but the two bounds are unequal for all intermediate values of R. In this correspondence we present a new upper bound for $\bar{D}(R)$:

$$1 - 2D(R) \ge \sup \{(1 - 2\alpha)^{R/(H_2(\alpha)+R-1)}: 0 \le \alpha \le \tfrac{1}{2}, H_2(\alpha) > 1 - R\} \quad (1.1)$$

which is smaller than Elias' bound for $R < 0.0509142$. A computer-generated graph of the three bounds is given in Fig. 1. More interesting than this small improvement, perhaps, is the fact that the new bound has the same derivative ($-\infty$) as the Gilbert bound at $R = 0$. (Elias' bound has derivative $-\ln 2$ at $R = 0$.) This fact supports the popular conjecture that $D(R) = f(1 - R)$. (Elias' bound has the same derivative (0) as Gilbert's at $R = 1$; our bound has derivative $-\infty$ at $R = 1$.) The new bound is based upon a nonlinear version of the MacWilliams identities, which we describe in Section II. The derivation of the new bound occupies Sections III and IV.

II. Nonlinear MacWilliams Identities

Let C be a binary block code of length n with $M = 2^{Rn}$ codewords. For each $i = 0,1,\cdots,n$, let N_i denote the number of ordered pairs of codewords (c,d) from C such that the Hamming distance between c and d is i. Define $a_i = N_i/M$. It follows that

$$a_0 = 1 \quad (2.1)$$

$$\sum_{i=0}^{n} a_i = M. \quad (2.2)$$

Let x be an indeterminate and define b_j, $j = 0,1,\cdots,n$ by the polynomial equation

$$\frac{1}{M}\sum_{i=0}^{n} a_i(1 - x)^i(1 + x)^{n-i} = \sum_{j=0}^{n} b_j x^j. \quad (2.3)$$

In 1963 MacWilliams showed that for a linear code, b_j = number of words of Hamming weight j in the dual code of C. In general, b_j appears to have no natural combinatorial significance, but it has recently been established (e.g., Delsarte [2]) that in any event $b_j \ge 0$, for all j. Everything depends on this innocent appearing result, and thus we have included a proof of it in the Appendix.

If x is replaced by $(1 - Z)/(1 + Z)$ in (2.3), a simple algebraic manipulation yields

$$\frac{1}{M}\sum_{i=0}^{n} a_i Z^i = \sum_{j=0}^{n} b_j \left(\frac{1 + Z}{2}\right)^{n-j}\left(\frac{1 - Z}{2}\right)^{j}. \quad (2.4)$$

Following Berlekamp [1, pp. 404–405], we multiply (2.4) by e^{nx}, set $Z = e^{-2x}$, take rth derivatives, and at $x = 0$

$$\frac{1}{M}\sum_{i=0}^{n} a_i(n - 2i)^r = \sum_{j=0}^{n} b_j F_r^{(j)}(n) \quad (2.5)$$

where $F_r^{(j)}(n) = d^r(\cosh^{n-j}(x) \sinh^j(x))/dx^r|_{x=0}$. The equations (2.5) are known as the Pless identities, since they were discovered for linear codes by Pless in 1963. It is easy to see that $F_r^{(j)}(n) \ge 0$, for all (r,j,n) and that $F_r^{(0)}(n) = 2^{-n}\sum_k \binom{n}{k} \cdot (n - 2k)^r$. Applying the result $b_j \ge 0$ cited earlier, we arrive at the main result of this section (note that $b_0 = 1$).

Theorem A: For $r = 0,1,2,\cdots$,

$$\frac{1}{M}\sum_{i=0}^{n} a_i(n - 2i)^r \ge \frac{1}{2^n}\sum_{k=0}^{n}\binom{n}{k}(n - 2k)^r = F_r^{(0)}(n).$$

(It can be shown that equality holds in Theorem A, for all r, if and only if C contains all 2^n binary codewords of length n.)

III. The New Bound

Applying the binomial theorem

$$(n + 1 - 2i)^r = \sum_{s=0}^{r}\binom{r}{s} n - 2i)^s$$

to Theorem A, we obtain

$$\frac{1}{M}\sum_{i=0}^{n}(n + 1 - 2i)^r a_i \ge \sum_{s=0}^{r}\binom{r}{s} F_s^{(0)}(n). \quad (3.1)$$

Assume r is odd (replace it with $2r + 1$), and bound the right side of (3.1) from below with its penultimate term

$$\frac{1}{M}\sum_{i=0}^{n}(n + 1 - 2i)^{2r+1}a_i \ge \frac{2r + 1}{2^n}\sum_{k=0}^{n}\binom{n}{k}(n - 2k)^{2r}. \quad (3.2)$$

Next, let d denote the minimum distance of the code C, i.e., the smallest positive j such that $a_j > 0$; $a_i(n + 1 - 2i)^{2r+1} \le a_i(n + 1 - 2d)^{2r+1}$, for all $i > 0$. Then (3.2) combined with (2.1) and (2.2) yields

$$\left(1 - \frac{2d}{n + 1}\right) \ge \left\{\frac{2r + 1}{2^n(n + 1)}\sum_{k=0}^{n}\binom{n}{k}\left(\frac{n - 2k}{n + 1}\right)^{2r} - 2^{-Rn}\right\}^{1/(2r+1)} = (s(n,r) - 2^{-Rn})^{1/(2r+1)} = e_r(n,R) \quad \text{definitions.} \quad (3.3)$$

If

$$E(R) = \lim_{n\to\infty}\frac{1}{n}(\sup_{r\ge 0} e_r(n,R))$$

then (3.3) implies that $\bar{D}(R) \le \frac{1}{2}(1 - E(R))$. The remainder of this correspondence is devoted to showing that the function appearing on the right side of (1.1) is a lower bound to $E(R)$. (Closer analysis shows that $E(R)$ is actually equal to the right side of (1.1); our bound on $\bar{D}(R)$ could not be improved even by retaining all of the terms on the right side of (3.1).)

IV. Estimation of $E(R)$

It follows from the definition (3.3) of $s(n,r)$ that

$$s(n,r) > \frac{2r+1}{2^n(n+1)}\binom{n}{k}\left(\frac{n-2k}{n+1}\right)^{2r}, \qquad \text{for all } k. \tag{4.1}$$

Let us define $\alpha = k/n$, $\beta = 2r/n$, and assume α and β are fixed and satisfy $0 < \alpha < \frac{1}{2}$, $\beta > 0$. Then (4.1) implies

$$\frac{1}{n}\log_2 s(n,r) > -1 + H_2(\alpha) + \beta \log_2 (1-2\alpha) + o(n), \qquad (0 < \alpha < \tfrac{1}{2},\ \beta > 0). \tag{4.2}$$

The estimate $1/n \log_2 \binom{n}{\alpha n} = H_2(\alpha) + o(n)$ follows from Stirling's approximation $\log n! = (n + \frac{1}{2}) \log n - n \log e + o(n)$.) If α and β are chosen so that $-1 + H_2(\alpha) + \beta \log_2 (1-2\alpha) > -R$, then $s(n,r)$ dominates 2^{-Rn}; hence $E(R) \geq \lim_{n\to\infty} s(n,r)^{1/(2r+1)}$, i.e.,

$$E(R) \geq 2^{(-1+H_2(\alpha))/\beta}(1-2\alpha),$$
$$(0 < \alpha < \tfrac{1}{2},\ \beta > 0,\ -1 + H_2(\alpha) + \beta \log_2 (1-2\alpha) > -R). \tag{4.3}$$

For fixed α, the right side of (4.3) is an increasing function of β, and thus the bound is maximized when β is as large as possible, subject to the given constraints. This largest possible β, $(1 - H_2(\alpha) - R)/\log_2 (1-2\alpha)$, is positive if $1 - H_2(\alpha) - R < 0$. Replacing β with this value, we obtain

$$E(R) \geq (1-2\alpha)^{R/(H_2(\alpha)+R-1)}, \qquad (0 < \alpha < \tfrac{1}{2},\ H_2(\alpha) > 1 - R). \tag{4.4}$$

This is the bound which was promised in Section I. We conclude with a proof that $E'(0) = \infty$; this implies that $\bar{D}'(0) = -\infty$ and that the new bound is definitely less than Elias' for sufficiently small R.

Lemma: $E'(0) = \infty$.

Proof: The bound (4.4) gives $E(0) \geq 0$, and the fact $D(0) = \frac{1}{2}$ cited earlier implies $E(0) \leq 0$. Thus $E(0) = 0$ and $E'(0) = \lim_{R\to 0} E(R)/R$. For each $R > 0$, define $\alpha < \frac{1}{2}$ by $H_2(\alpha) = 1 - R/3$. Then by (4.4) $E(R) \geq (1-2\alpha)^{3/2}$, but it is easily verified that $\lim_{R\to 0} (1-2\alpha)^{3/2}/R = \infty$.

Appendix

The proof that $b_j \geq 0$ follows. The result needed in the paper is the case $q = 2$ of the theorem proved in this appendix. Let $S = \{0, 1, \cdots, q-1\}$ be the cyclic group of the integers modulo q, q being an arbitrary integer ≥ 2. Let ζ be a primitive complex qth root of unity, and for $s, t \in S$, define $(s,t) = \zeta^{st}$.

Lemma:

$$\sum_{s\in S}(s,t) = \begin{cases} 0, & \text{if } t \neq 0 \\ q, & \text{if } t = 0. \end{cases}$$

Proof: If $t = 0$, the assertion of the lemma is trivial. If $t \neq 0$, ζ^t is a zero of $x^{q-1} + x^{q-2} + \cdots + 1$ and

$$\sum_{s\in S}(s,t) = \sum_{s=0}^{q-1}\zeta^{st} = 0.$$

For any integer n denote by S^n the group of n-tuples $(s_1, s_2, \cdots, s_n)$ $s_i \in S$, addition componentwise, and extend the definition $(\cdot,\cdot)$ to S^n by

$$(u,v) = \prod_{i=0}^{n}(u_i, v_i).$$

If $a = (a_1, a_2, \cdots)$ is any finite list from S, define $w(a)$ as the number of nonzero entries in a. Then for $u, v \in S^n$, $w(u - v)$ is the Hamming distance between u and v.

Let C be code of length n and rate R over S, i.e., a subset of S^n of cardinality $M = q^{Rn}$. For each $i = 0,1,\cdots,n$, let a_i be defined by

$$a_i = \frac{1}{M}|\{(u,v): u, v \in C,\ w(u - v) = i\}|.$$

Theorem: Let x be an indeterminate, and define real numbers b_j by the polynomial equation

$$\frac{1}{M}\sum_{i=0}^{n} a_i(1-x)^i(1 + (q-1)x)^{n-i} = \sum_{j=0}^{n} b_j x^j.$$

Then $b_j \geq 0$, for all j.

Proof: First we show that

$$\sum_{u\in S^n} x^{w(u)}(u,v) = (1-x)^{w(v)}(1 + (q-1)x)^{n-w(v)}. \tag{A.1}$$

To prove (A.1)

$$\sum_{u\in S^n} x^{w(u)}(u,v) = \sum_{u_1=0}^{q-1}\cdots\sum_{u_n=0}^{q-1} x^{w(u_1)+\cdots+w(u_n)}(u_1,v_1)\cdots(u_n,v_n)$$
$$= \prod_{i=1}^{n}\sum_{s=0}^{q-1} x^{w(s)}(s,v_i).$$

This last sum is $(1 + (q-1)x)$, if $v_i = 0$ and is

$$1 + x\sum_{s=1}^{q-1}(s,v_i) = 1 - x$$

if $v_i \neq 0$, by the lemma. This establishes (A.1).

To prove the theorem, sum both sides of (A.1) over the M^2 vectors $c - d$, $c, d \in C$:

$$\sum_{c,d\in C}\sum_{u\in S^n} x^{w(u)}(u, c-d)$$
$$= \sum_{c,d\in C}(1-x)^{w(c-d)}(1 + (q-1)x)^{n-w(c-d)}$$
$$= M\sum_{i=0}^{n} a_i(1-x)^i(1 + (q-1)x)^{n-i}$$

by the definition of a_i. On the other hand

$$\sum_{u\in S^n} x^{w(u)}\sum_{c,d\in C}(u, c-d) = \sum_{u\in S^n} x^{w(u)}\sum_{c\in C}(u,c)\sum_{d\in C}(u,-d)$$
$$= \sum_{u\in S^n} x^{w(u)}\left|\sum_{c\in C}(u,c)\right|^2$$

since $(u,-d) = \overline{(u,d)}$. This completes the proof, and, in fact, shows that

$$b_j = \frac{1}{M^2}\sum_{\substack{u\in S^n \\ |u|=j}}\left|\sum_{c\in C}(u,c)\right|^2.$$

References

[1] E. R. Berlekamp, *Algebraic Coding Theory*. New York: McGraw-Hill, 1968.
[2] P. Delsarte, "Bounds for unrestricted codes, by linear programming," *Philips Res. Rep.*, vol. 27, pp. 272–289, 1972.